

This week:

1. This week I take four days, from Monday to Thursday, to improve our system and make a video. Video including three cases, lost phone, analyze traffic jam and analyze social network of citizens from village.
2. Revise paper. Mainly revise cover letter and case study. The rest of the paper I have written in Chinese, It will be finished in next two days and I will send an e-mail.
3. Review a paper
4. Read paper

- Patterns and Sequences: Interactive Exploration of Clickstreams to Understand Common Visitor Paths

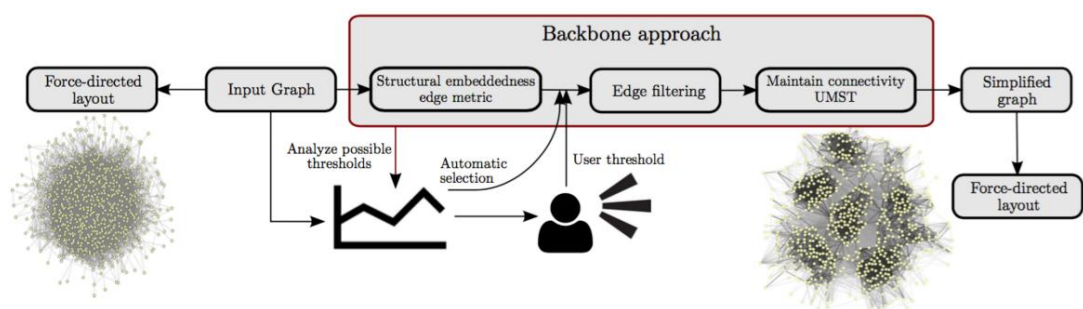
本文研究的是针对于某一网站的用户访问点击事件的时序数据，文章如果直接对数据进行可视化，则难以发现有效地行为模式。然而如果仅仅对数据进行简单的聚类，那么则会因为数据聚类过多而导致结果复杂，难以挖掘有用的信息。本文采用了模式的消除算法。如果两个模式的原始数据距离比较小，则合并这两个模式。

文章中的系统采用了两个界面，模式界面和数据序列界面，模式视图主要展示不同的点击模式，用颜色编码。由上到下表示时间。每一个模式的序列长度则用矩形的长度表示。当用户选择一个模式的时候，该模式的序列详细视图则会展现在右边的矩阵时间视图中。提供用户进一步的探索



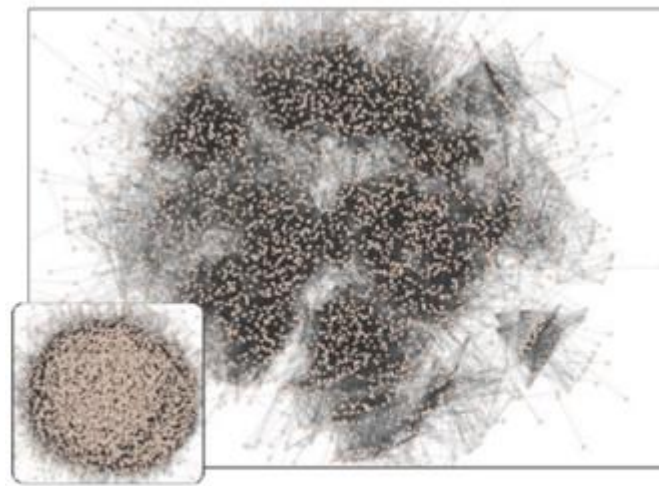
- Adaptive Disentanglement based on Local Clustering in Small-World Network Visualization

由于之后的本体论和城市数据相结合，估计会用到节点连接图的简化渲染，所以看了这篇文章。本工作采用了对节点连接图预处理的方法，自动选择过滤的阈值阈值，不显示图中不重要的边，只显示重要的图结构，使得节点图更加清晰。



当节点连接图画出时，将计算该图形的节点是否与其相关联的节点联系位置密切。从

而决定过滤节点边的阈值，当去掉节点中的边的时候进行力引导使得节点之间更加分散。通过本文的算法，应用到我们的通话记录形成的社交网络中应该会十分有效。因为我们可能只对那些重要的关键性的通话感兴趣，并可以把保留的边（通话）作为进一步分析组聚类的有效数据，进行加权处理。



(c) Rice31,  $|V| = 4k$ ,  $|E| = 97k$

- 时空轨迹聚类

主要浏览了很多关于轨迹聚类的方法和文章，方便下一个项目的社交网络聚类的构思。记录如下

1. 最常见的方法是轨迹间欧氏距离。但是传统的欧式轨迹计算方法都不能应用于采样率不同或者尺度不同的轨迹数据

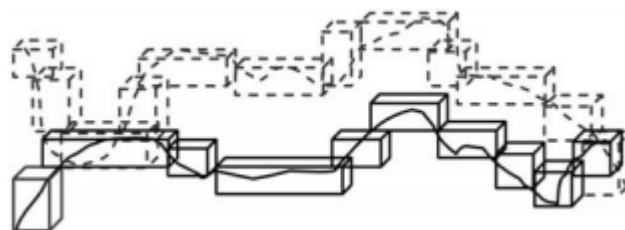
《Shape-based similarity query for trajectory of mobile objects》先将轨迹分段线性表示，然后内插重采样，再计算轨迹间欧氏距离，这样处理能够将采样率不同的轨迹数据进行比较。

《Indexing large human-motion database》先对轨迹进行全局缩放再计算轨迹间欧式距离，该方法能够有效解决尺度不同的问题

2. 最小外包矩形距离

《Similarity search for multidimensional data sequences》它首先将整条轨迹划分成一些相对平滑的轨迹区间，再将每条子轨迹用其最小外包矩形表示，各对外包矩形间的距离加权平均作为整体轨迹间的距离。

《Incremental clustering of mobile objects》将最小外包矩形重叠部分的大小作为整条轨迹相似性度量。



3. 全区间变换对应相似的聚类方法

《Rotation invariant distance measures for trajectories》时空轨迹的时间维可以局部拉伸和缩放，只需要保证轨迹记录点的时间顺序，而不需要在一一对应

的时刻上进行比较。DTW(Dynamic Time Warping)是典型代表

Next week:

Finish VAUD before Tuesday.

Read 5 paper.